

THE UNIVERSITY
OF ILLINOIS
LIBRARY

370
IL6e
no.25-37

~~EDUCATION~~

Return this book on or before the **Latest Date** stamped below. A charge is made on all overdue books.

University of Illinois Library

JUN - 1 1954

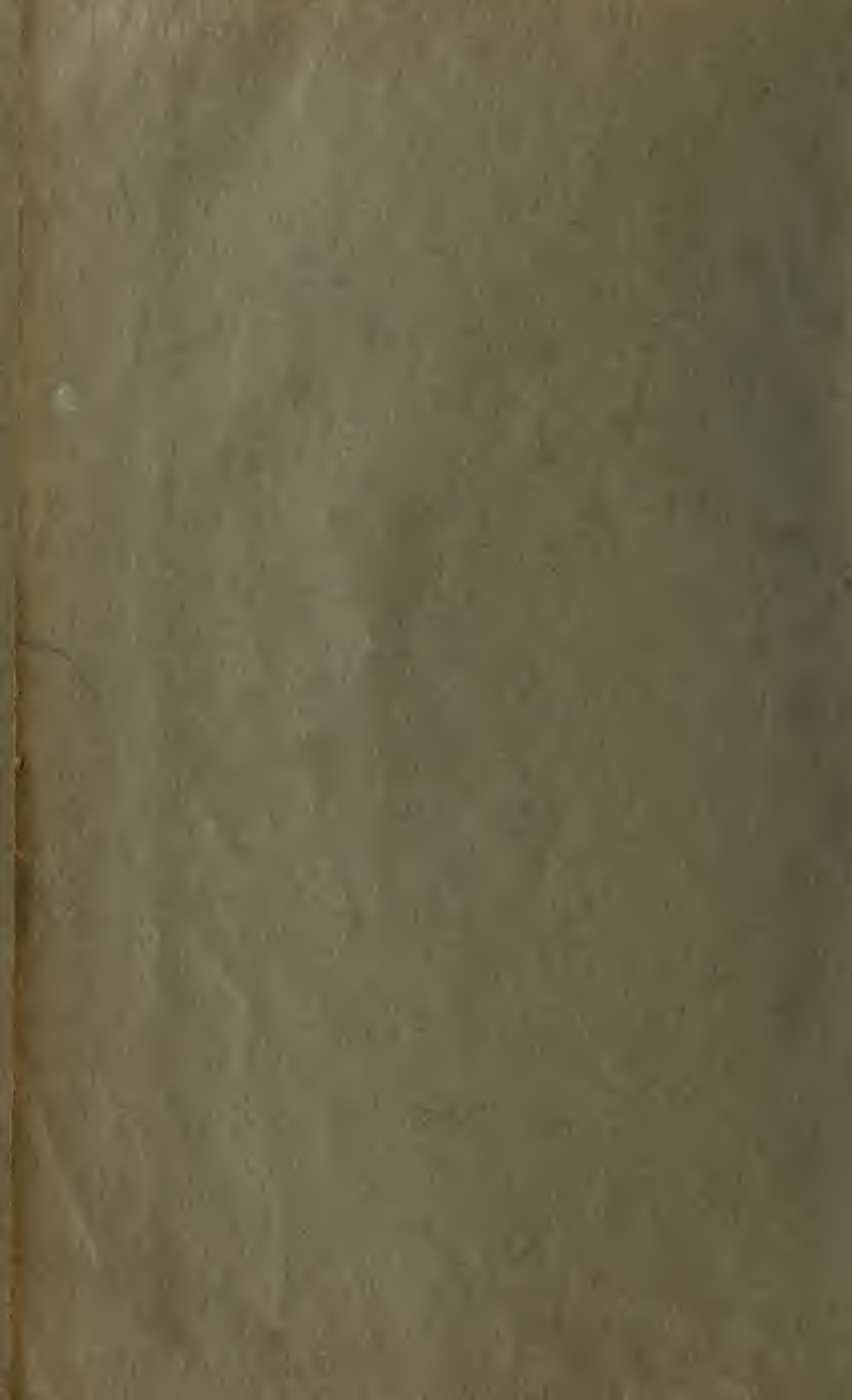
AUG 12 1954

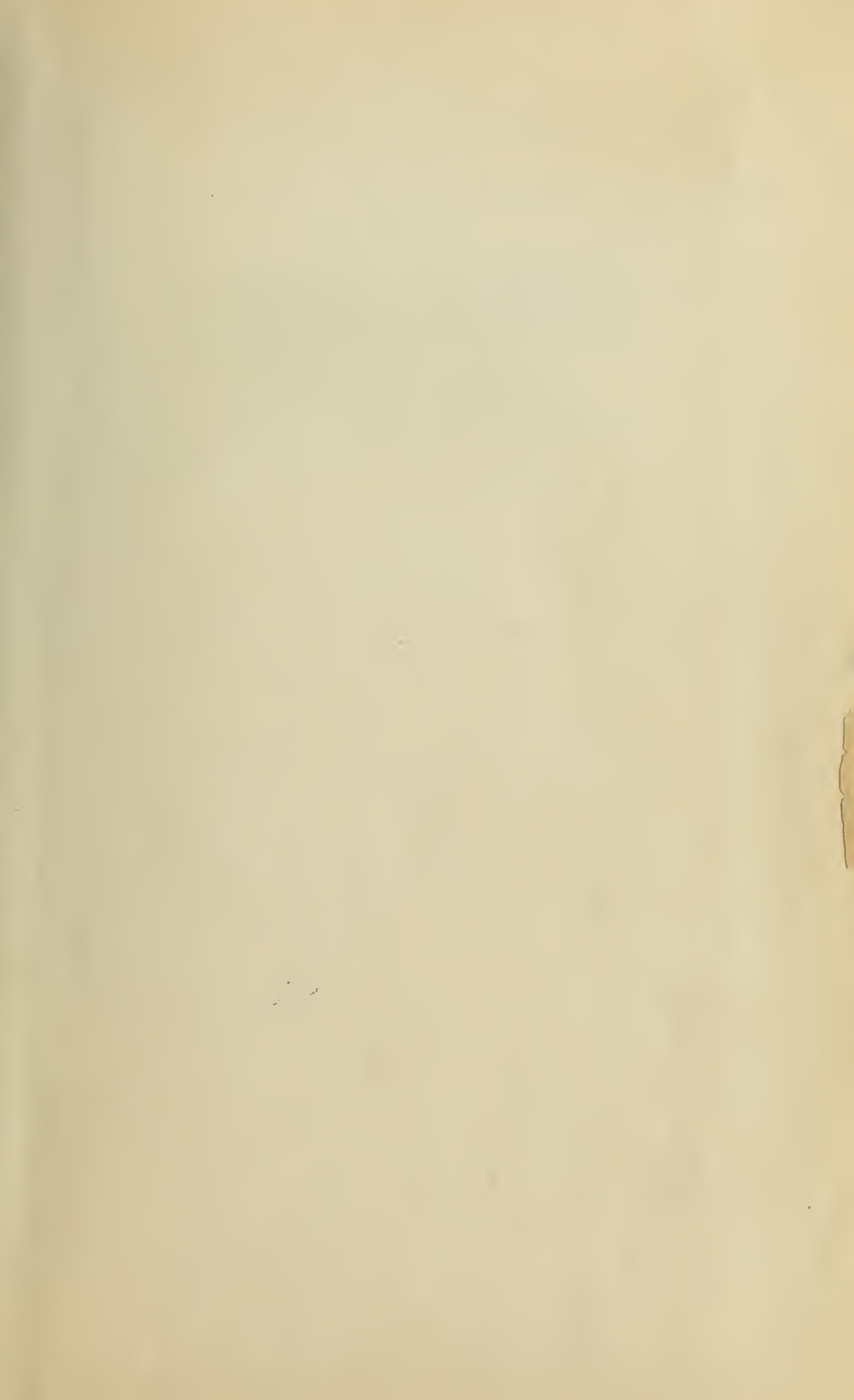
7/20/53


JUN - 9 1955

DEC - 6 1965

MAR 15 1981







Digitized by the Internet Archive
in 2012 with funding from
University of Illinois Urbana-Champaign

SECRET

0.25

UNIVERSITY OF ILLINOIS BULLETIN

ISSUED WEEKLY

Vol. XXI

JANUARY 28, 1924

No. 22

[Entered as second-class matter December 11, 1912, at the post office at Urbana, Illinois, under the Act of August 24, 1912. Acceptance for mailing at the special rate of postage provided for in section 1103, Act of October 3, 1917, authorized July 31, 1918.]

EDUCATIONAL RESEARCH CIRCULAR NO. 25

BUREAU OF EDUCATIONAL RESEARCH
COLLEGE OF EDUCATION

MEASURING TEACHING EFFICIENCY

By

WALTER S. MONROE

Director, Bureau of Educational Research

and

JOHN A. CLARK

Assistant, Bureau of Educational Research



THE UNIVERSITY OF THE
JAN 25 1927

UNIVERSITY OF ILLINOIS

PUBLISHED BY THE UNIVERSITY OF ILLINOIS
URBANA

370
I. 66e
no. 25-37

LIBRARY
DEPARTMENT

Measuring Teaching Efficiency

Until recently teachers were rated by what has been called the "general impression method." Under this method the judgments of a supervisor were controlled neither by an outline nor by other specifications. Obviously different supervisors would vary widely in the judgments expressed with references to the same teacher. Furthermore, a given supervisor was likely to judge different teachers on different bases or to be influenced by some minor detail either favorable or unfavorable. For obvious reasons this method of measuring the efficiency of individual teachers is unsatisfactory. Beginning about 1910, a number of investigators attempted to devise procedures that would yield measurements of teaching efficiency which had a definite meaning and were more accurate than could be obtained by the "general impression method." The methods which have been proposed will be considered under the following items:

- I. Score cards.
- II. Man-to-man comparison scale.
- III. Measurement of teaching efficiency by means of standardized tests.

I. MEASUREMENT OF TEACHING EFFICIENCY BY MEANS OF SCORE CARDS

Beginning of teacher score cards. The earliest investigations, among which may be listed the work of Book (1905), Kratz (1907), Ruediger and Strayer (1910), Boyce (1912), Littler (1914) and Moses (1914), were attempts to analyze teaching efficiency and to identify the essential traits or characteristics of successful teachers. The efforts of these early investigators have not met with complete approval and a number of more recent attempts to formulate a list of characteristics which would determine the essential traits of successful teachers have been made. Among these may be noted Clapp (1915), Buellesfield (1915), Mead and Holley (1916), Johnston (1916), Osborn (1920), and Knight (1922).

Using the results of the earlier studies as a basis, Elliott in 1910 formulated a score card for measuring teaching efficiency. This card consisted of a list of forty-two traits which were considered essential to successful teaching. The teacher was judged with reference to

each of these forty-two traits. The sum of these judgments, which were to be expressed in numerical terms, was taken as a measure of the teacher's efficiency. Since 1910 a number of score cards have been formulated by other persons. Although there are many points of similarity in these scales they differ in certain details. Those by Elliott (1910), the New York Bureau of Municipal Research (1915), Boyce (1915), Landsittel (1917), Rugg (1920), Connor (1920), Kent (1920), Maddock (1922), and Carrigan (1922), appear to be typical of the differences in structure.

Functions of score cards. Score cards for measuring teaching efficiency fulfill two functions. They may be used by the superintendent for administrative purposes such as a basis for determining reemployment, promotions and salary increases or they may serve as a means for improving the teachers in service. When used for administrative purposes the rating should be made by the superintendent, the principal or other supervisory officials. When the function of the score card is that of teacher improvement it has been recommended that each teacher be asked to rate himself. This process of rating, however, is considered self-analysis rather than measurement. In some cases a preliminary rating has been made both by the principal or special supervisor and by the teacher and a final rating given by the superintendent after he has compared the results of the two ratings and has conferred with the individual teachers regarding any special characteristics of strength or of weakness.

Representative score cards described. 1. Elliott's Score Card for Measuring the Merits of Teachers. This score card has two general divisions: (I) Individual Efficiency, and (II) Directed Efficiency. Under the former there are eight main headings: (1) Physical Efficiency, (2) Moral-Native Efficiency, (3) Administrative Efficiency, (4) Dynamic Efficiency, (5) Projected Efficiency, (6) Achieved Efficiency, and (7) Social Efficiency. A number of subordinate traits are also given. For example, under Administrative Efficiency the following are listed: (1) Regularity at post of duty, (2) Initiative, resourcefulness, (3) Promptness and accuracy, (4) Executive capacity, (5) Economy of time and property, (6) Cooperation with associates and supervisors. In the second division there is only one heading, Supervisory Efficiency.

2. Score card by New York Bureau of Municipal Research. The score card devised by the New York Bureau of Municipal Research (1915) differs in structure from that by Elliott. It appears to

have been designed for the purpose of securing a qualitative description of the teacher rather than a numerical rating. A check mark is to be placed opposite descriptive terms in order to note their presence or degree. Two sections of the card are reproduced below:

I. PERSONALITY OF TEACHER (check \checkmark)

(a) Teacher appears to be
 vigorous.....weak.....poised.....nervous.....
 neat.....slovenly.....at ease.....embarrassed.....

II. TEACHING ABILITY AS SHOWN BY (check \checkmark)

(a) Extent to which teacher's questions are	Not at all	Slight	Medium	Notable	Evidence and Remarks
(1) thought-provoking.....					
(2) calling for facts.....					
(3) suggesting the answer.....					
(4) answered by "yes" or "no".....					
(5) irrelevant.....					
(6) not definite—vague.....					

3. Rugg's Scale for Rating Teachers in Service. Rugg's Rating Scale for judging teachers in service was designed to be used primarily by the teacher in analyzing and rating himself. Its structure is unique. It consists of over fifty questions which are grouped under five heads: (I) Skill in Teaching; (II) Skill in the Mechanics of Managing a Class; (III) Team Work Qualities; (IV) Qualities of Growth and Keeping Up-to-Date; (V) Personal and Social Qualities; and which are to be answered in terms of "low," "average," or "high." The questions under the first heading are reproduced.

To what extent—

Does he know the subject matter of his own and related fields:

1. In subjects like history, geography, etc., does he make effective use of material outside the text book.....
2. Does he relate lessons to material in other fields and use illustrations outside his own subject (e. g., math. and science).....

Does he select subject-matter effectively for class reading and discussion.....

Are his aims of teaching clearly defined.....

Does he give evidence of having:

1. Formulated clearly his aims of teaching, as shown by his written statement of aims and outcomes.....
2. Planned his lessons specifically to carry these out.....
3. Distinguished clearly between (a) "formal skill" (either in manual or academic subjects), (b) "information" and (c) "problem solving" as proper outcomes from his class work.....

4. Given pupils clear ideas of the purpose of lessons.....
- Is he skillful in conducting the class discussion.....
 - a. Resourcefulness in organizing a discussion and in "thinking on his feet".....
 1. Is he fertile and quick in taking advantage of pupils' questions.....
 2. Are his questions systematically planned, yet spontaneously given.....
 3. Does he express himself clearly.....
 - b. Skill in conducting "drill" exercises.....
 1. Does he make use of economical, "timed," drill-devices (such as Curtis' Practice Exercises, etc).....
 2. Does he properly subordinate drill to clear exposition; that is, keep a proper balance between drill and "development".....
 - c. Ability to "develop" new phases of the work.....
 1. Are lessons well related to previous ones.....
 2. Is material "organized".....
 3. Do lessons show the use of material in the solution of present or future problems:
 - a. In his subject.....
 - b. Outside his subject.....
 - d. Ability to secure class participation in the recitation.....
 1. Do all pupils in the class take part in the discussion.....
 2. Do all the pupils question each other and conduct the class independently of his formal direction.....
 - e. Skill in making the assignment.....
 1. Was it an attempt to teach pupils how to study the lesson.....
 2. Was it more than mere formal announcement of the number of pages in the text, etc.....
 3. Are its scope and purpose clearly recognized by pupils.....
- Has he insight into "how children learn".....
 1. Does he keep the discussion within the pupils' comprehension.....
 2. Does he endeavor to discover pupils' difficulties by keeping records of errors and studying these.....
 3. Does he adapt discussion to individual differences in pupils.....

✦ **Distinctive characteristics of other score cards.** Score cards for rating teachers have varied widely with reference to the number of traits upon which teachers are to be rated. Boyce included forty-five; Landsittel thirty-four; and Maddock reduced the number to eight general traits. Carrigan observed one criterion which does not appear to have been recognized by other makers of score cards. She rejected those traits regarding which a supervisor could not be expected to express a judgment as the result of a single visit of one period to the teacher's classroom. Among such traits are health, disposition of teacher, and cooperation. Kent has criticized score cards by pointing out that the achievements of pupils are given relatively little weight, and proposes a scheme of rating in which the achievements of pupils are considered as the most important single division. Connor, by giving explicit definitions of the qualities to be appraised, has attempted to make ratings by means of score cards more objective. He also gives a prominent place to the achievements of pupils.

Devices for assigning ratings to traits. A variety of devices have been proposed for assigning ratings to the traits enumerated on the score cards. Elliott's gave the number of points to be given for perfection in each of the traits. A supervisor was then instructed to deduct from this maximum. For example, if the maximum credit allowed a trait was 10 the deductions for deficiencies were to be as follows: very slight, 2; slight, 4; marked, 6; very marked, 7; extreme, 8. In Boyce's score card a quality rating was given for each trait by placing a check mark in the appropriate column. These qualitative ratings, however, could easily be translated into numerical equivalents if such were desired. The score card of the New York Bureau of Municipal Research did not provide for numerical ratings. The same is true of the card by Rugg, although it would be possible to devise a procedure for translating the qualitative ratings into numerical ones.

Score cards do not yield accurate measures of teaching efficiency. In judging the worth of any measuring instrument attention is centered primarily upon two questions: first, does it measure the trait or traits which it claims to measure; second, how accurate are the measurements which it yields. Score cards have been shown defective in both respects, but because the gross inaccuracies in the measures yielded are sufficient to disqualify them as a practical device for the measurement of teaching efficiency, it is not necessary in this circular to consider also their limitations with respect to overlapping of traits or to the particular traits enumerated.

Rugg¹ states that he obtained coefficients of reliability for the Elliott Scale for Measuring the Merits of Teachers which closely approximated 0. "Practically no correlation exceeded .2." He also states that the probable error of measurement for such scales as Elliott's, Boyce's, Beatty's and Hill's is approximately 10 points on a scale of 100 points. He further asserts that we should discard ordinary rating scales for measuring teaching efficiency. "We cannot justify wasting the time of our school administrators and deluding our teachers with fictitious 'ratings' and 'marks.' Even on one of the so-called 'standardized' point rating schemes *single* rating has little or no scientific validity." In another place² he claims that "the un-

¹Rugg, H. O. "Is the rating of human character practical?" *Journal of Educational Psychology*, 12:426, November, 1921.

²Rugg, H. O. "Self-improvement of teachers through self-rating: a new scale for rating teachers' efficiency," *The Elementary School Journal*, 20:683, May, 1920.

reliability of current typical ratings of teachers by principals is so great that it is almost valueless."

After an extensive inquiry into the qualities related to success in teaching and their measurement, Knight³ concludes "that in judging particular traits general estimate influences the particular estimate to such a degree that judgments of particular traits are in themselves of little practical use." A supervisor's rating of a teacher in some particular trait is "a defense of his general estimate of that teacher as well as a rating of the trait under consideration." Incidentally, it is significant to note that Knight also concludes that "the general factor of interest in one's work becomes the dominant factor in determining one's success in teaching."

Score cards useful as a means of self-improvement. Although the conclusion that score cards are unsatisfactory as a means of measuring teaching efficiency cannot be avoided, it does not necessarily follow that they have no value. As pointed out in the beginning, score cards have been thought of as fulfilling two functions. In addition to their use as instruments for securing a measure of teaching efficiency many supervisors have found them very helpful as a means of improving teachers in service. In fact, some of the more recent score cards, for example the one by Rugg, have been designed with this function specifically in mind.

II. MAN-TO-MAN COMPARISON SCALES

Origin. A different type of instrument for measuring teacher efficiency was originated by Walter Dill Scott at the Carnegie Institute of Technology in 1917. The essential feature of this plan, which is called a Man-to-Man Comparison Scale, was that the supervisor or any person using the scale made one of his own which consisted of real persons intimately known to him. These scale persons, five in number, are chosen so that they represent degrees of the traits or general qualities from the poorest to the best. In making such a scale to measure teacher efficiency, one is directed to select "the best teacher you have ever known" for the highest step of the scale. For the lowest step of the scale "the poorest teacher you have ever known" is to be selected. Other teachers are chosen to represent "average," "better than average," and "poorer than average." This scale is to be used in very much the same way as one in handwriting

³Knight, F. B. "Qualities related to success in teaching." Teachers College Contributions to Education No. 120. New York: Teachers College, Columbia University, 1922. 67 p.

or composition. In measuring a given phase of teacher efficiency, the teacher under consideration is compared with the scale teacher. The scale value of the scale teacher which he is judged to resemble most nearly is taken as a measure of this phase of his teaching efficiency.

This man-to-man comparison scale was first used by Scott for rating employees in industry. When the United States entered the World War in 1917, Scott and a number of other eminent psychologists were called into service for the purpose of devising means for rating prospective as well as commissioned officers. The man-to-man comparison scale was decided upon as the device most likely to give satisfactory results. The technique for the construction of such a scale and its application was worked out. In September, 1918, Rugg was invited to make a study of this scale, particularly of its reliability.⁴

How to make a Man-to-Man Comparison Scale for measuring teaching efficiency. A unique feature of the man-to-man comparison scale is that each person makes his own scale. In doing this it is highly important that he exercise care in selecting scale teachers who will accurately represent the degree of the quality for which they are chosen. It is obvious that anyone who does not have a reasonably wide acquaintance with teachers will be seriously handicapped in making a scale. The following procedure is recommended.

1. The first step is to decide upon the rubrics of qualities or traits which are to be measured. It is probably wise to recognize from three to five rubrics of qualities rather than to make one general scale including all qualities. Rugg⁵ has suggested the following divisions: (1) Skill in teaching, (2) Skill in mechanics of managing a class, (3) Team work qualities, (4) Quality of growth and keeping up-to-date, (5) Personal and social qualities.

2. For each rubric of qualities a separate scale is to be made. The first step in its construction is to write down the names of at least twenty-five teachers whom you know well with respect to the first group of traits being considered. This list should include teachers representing various degrees of excellence. It is highly important that the poorest teacher and also the best teacher whom you have ever known be included. This procedure is to be repeated for each rubric. It is likely that the names of certain teachers will appear in two or more lists.

⁴For the result of Rugg's investigation see page 11.

⁵Rugg, H. O. "Self-improvement of teachers through self-rating: a new scale for rating teachers' efficiency." *Elementary School Journal*, 20:670-84, May, 1920.

3. Arrange the teachers in each list in order of merit, placing the best at the top and the poorest at the bottom. In doing this consider only the traits included in the rubric under consideration. For example, in arranging the teachers in "skill in managing a class" no consideration should be given to their "qualities of growth and keeping up-to-date." This ranking of teachers is the most important as well as the most difficult step in the formation of the scale, and should be made as accurate as possible.

4. From each list select five teachers who will be satisfactory representatives of the following degrees of excellence: (1) "the best teacher ever known," (2) "better than average teacher," (3) "average teacher," (4) "poorer than average teacher," and (5) "the poorest teacher ever known." It is recommended that these five teachers be selected in the following order: "best," "poorest," "average," "better than average," "poorer than average."

The final scale consists of the list of five teachers for each rubric of qualities selected for measurement. When a supervisor has once made a scale it may be preserved and used year after year until there is good reason for revision. Such a scale should be a part of the equipment of each supervisor.

In a large city school system where there are several supervisors engaged in rating teachers it will be helpful to have them work together in preparing their man-to-man comparison scales. By such cooperation greater uniformity in the measurements will be secured even though the completed scales probably will not be composed of the same teachers.

Method of rating teachers by means of a man-to-man comparison scale. In rating teachers with a man-to-man comparison scale only one rubric of qualities is considered at a time. Numerical values are assigned to each of the scale teachers. The following have been proposed: 38, 30, 22, 14, 6. If a particular teacher when rated with respect to "skill in teaching" is judged to be equivalent to the best teacher the supervisor has ever known he would receive a numerical rating of 38. On the other hand if he is judged to be equivalent to the "poorer than average" he would receive a rating of 14. In rating a teacher with reference to one rubric of qualities no consideration should be given to his other qualities or to the other qualities of the scale teacher. The total rating of a teacher is obtained by adding together the numerical ratings on each of the rubrics of qualities.

The reliability of ratings by means of a man-to-man comparison scale. No study has been made as yet of the reliability of teacher ratings by means of a man-to-man comparison scale when used by superintendents or other supervisors. The reliability of such ratings can be inferred from the study which Rugg made of the use of this scale in the army. For two ratings of the same officers by different persons he states that the average differences between the two ratings were:

“For second lieutenant.....12.0 points
For first lieutenant.....21.7 points
For captain.....16.9 points.”

The maximum possible rating in all cases was 80 points. Rugg also states that “it was very improbable that an officer was located within even his proper ‘fifth’ of the entire scale in his ‘official’ rating,” and “the chances can not be more than four to one that any rating will be within fourteen points of the persons true rating.” The probable error of measurement is approximately seven points on the scale of eighty.

Man-to-man comparison scale versus score cards. It is clear from the information presented in the preceding discussions that neither score cards nor man-to-man comparison scales may be expected to yield highly accurate measures of teaching efficiency. Even under the most favorable conditions the probable error of measurement will be so large that serious limitations must be placed on the measures secured. It is, however, worth while to note that the measures yielded by the man-to-man comparison scale will ordinarily be more accurate than those secured by the usual score card.

III. MEASUREMENT OF TEACHING EFFICIENCY BY MEANS OF STANDARDIZED TESTS

Achievements of pupils an index of teaching efficiency. The proposal to measure the efficiency of a teacher by means of standardized tests is based upon the thesis that “a teacher’s merit is directly proportional to the changes which he engenders in his pupils.” His training, personality, initiative, health, skill in teaching, ability as a disciplinarian, etc., are significant only for the effects which they produce in the pupils. In other words, “By their fruits ye shall know them.”

Teacher activity versus pupil activity. Both of the preceding methods of measuring teaching efficiency have implied the opposite point of view, viz., that the teacher should be judged by his traits and activities rather than by the achievements of his pupils.⁶ In the use of practically all the score cards and man-to-man comparison scales, attention is focused primarily upon the activity of the teacher. In the use of standardized tests the attention is transferred to pupil activity and achievements. The first method implies that the things which the teacher does are ends in themselves; the second, that these things are merely means to an end.

Limitations of the measurement of teaching efficiency by means of standardized tests. The use of standardized tests for the measurement of teaching efficiency is limited by the lack of such tests for measuring all of the outcomes engendered by the teacher. Skills in the tool subjects, reading, arithmetic, hand-writing, spelling, and language form an important group of achievements, but the pupil is expected to acquire also other habits and knowledge and to form desirable ideals, tastes, interests and perspectives. These last constitute the "less tangible outcomes" of education, and for measuring them no satisfactory standardized tests have been devised. Hence, *all* of the results of teaching cannot as yet be measured. Moreover, when standardized tests are used systematically as instruments for measuring teaching efficiency, teachers and pupils very likely tend to emphasize the tool subjects which can be measured and to neglect the less tangible but equally important outcomes. Such test results cannot be considered true indications of teaching efficiency.

Certain imperfections of our present standardized tests constitute additional limitations. Variable errors of measurement tend to be neutralized in the average scores of a class, but constant errors cannot be eliminated in this way, and make measures of achievements erroneous indices of teaching efficiency.⁷

Since the engendering of skills is much less prominent in the work of the high school than in that of the elementary school it naturally follows that in the former fewer standardized tests are available for measuring teaching efficiency. In fact, except for algebra, foreign languages, typewriting, and stenography there are prac-

⁶Kent, R. A. "What should teacher rating schemes seek to measure?" *Journal of Educational Research*, 2:802-07, December, 1920.

⁷For a further discussion of teaching efficiency see, Monroe, Walter S. "The constant and variable errors of educational measurements." *University of Illinois Bulletin*, Vol. 21, No. 10. Bureau of Educational Research Bulletin No. 15. Urbana: University of Illinois, 1923.

tically no standardized tests for high-school use that may be considered satisfactory.

Standardized tests yield one index of teaching efficiency. In spite of their limitations it should be recognized that in the elementary school standardized tests, in that they measure specific habits or skills, do yield an important index of teaching efficiency. A superintendent or principal would be distinctly unwise in basing his measures of the merits of his teachers wholly upon the results of standardized tests, but he would also be unwise if he did not consider these results in making his final estimate of the worth of his teachers. Furthermore, the use of standardized tests serves to focus the attention upon pupil activity rather than upon teacher activity. Even in our attempts to measure teaching efficiency by other means it is probable that more valid results will be obtained if we consider the activity of the pupil rather than that of the teacher.

Quality of pupil material must be considered. In the use of standardized tests to measure teaching efficiency, it is necessary to take into account the quality of the pupil material as well as the achievements of the pupils. Cases have been reported in which teachers were grossly misjudged because they were working with a pupil group whose average intelligence was either very high or very low. Recently the achievement quotient (A. Q.) has been proposed as a device for expressing a measure of achievement in comparison with capacity to learn. Such a quotient is obtained by dividing a pupil's achievement by his general intelligence. Both measures must necessarily be expressed in terms of comparable units. A convenient method states the measures of achievement in terms of achievement age and the measures of general intelligence in terms of mental age. The quotient is expressed as a percent, the decimal point being omitted. A quotient of 100 means that a pupil's achievement is just equivalent to that of the average pupil of his mental age. Consequently an average or median achievement quotient for a class means that the teacher has been doing just average work. A median A. Q. of distinctly above 100 indicates superior teaching ability; on the other hand a median A. Q. distinctly below 100 indicates inferior teaching.

Even in the use of the achievement quotient as an index of teaching efficiency it is necessary to bear in mind the previous status of the class. The achievements of the group of pupils at any one

time is dependent in part upon their present teacher and in part upon their previous educational experience.

IV. SUGGESTED PLAN FOR THE MEASUREMENT OF TEACHING EFFICIENCY

The practical need for the measurement of teaching efficiency. Many superintendents and principals face the practical problem of securing a numerical rating of the efficiency of their teachers for the purpose of determining reemployment, promotions and increases in salary. Although none of the three methods for measuring teaching efficiency which have been considered in the preceding pages is satisfactory, it seems wise to suggest a plan to meet this practical need. The following procedure represents merely the opinion of the writers and is recognized as being imperfect. It is included in this circular because of several requests which have come to the Director of the Bureau of Educational Research for advice in regard to the rating of teachers.

Measurement of teaching efficiency should be based upon achievements of pupils. A teacher's academic and professional training, experience, intelligence, personal or social qualities, interest in teaching, and other traits are merely means to an end, namely, the engendering of achievements in school children. Thus the measure of a teacher's efficiency should be based upon the achievements which he engenders. In arriving at a measure of the efficiency of an operator of a machine only the output is considered. No attention is given to training and experience, interest in work, or other traits. The operator who obtains the greatest output is considered as the most efficient and the one whose production is low as inefficient. In the same way that teacher should be considered most efficient who engenders in his pupils the greatest growth in achievement and that teacher as least efficient who engenders the least growth.⁸ It should be noted that all the elements of growth must be measured; ideals, interests and attitudes must be considered as well as skills and knowledge.

Since we are limited in the measurement of school achievements and in their evaluation in terms of social worth, it is necessary that

⁸In evaluating this growth the social worth of the achievements must be considered. Certain achievements may be of little value when judged by their social usefulness. For example, the ability to spell words which are seldom, if ever, used in ordinary writing has little social value. Exceptionally high degrees of skill in performing the arithmetical operations have little general social value.

other factors be recognized in a practical plan for the measurement of teaching efficiency. In the plan proposed four such qualities are included: (1) personal and professional qualities, (2) general intelligence, (3) experience, (4) academic and professional training. As indicated above these traits are merely a means to an end but they sustain a fairly high positive correlation with the achievements of pupils.

Teaching experience and academic and professional training are more important in selecting teachers for employment than in measuring their efficiency after employment.⁹ They are included in this plan of rating because they permit of objective measurement. It should be noted that they are given relatively less weight than either of the other divisions of the plan.

I. ACHIEVEMENTS OF PUPILS

In measuring the achievements of pupils standardized tests should be used in so far as they are available and the achievement quotient (A. Q.) calculated. However, it will be necessary to supplement such measurement by means of written examinations and by teachers' estimates in the case of such outcomes as interest of pupils in school work, technique of study and ideals. There should be a distinct effort to secure a composite measure of all the outcomes of instruction of the teacher whose efficiency is being measured.

In judging the measures of achievement it is necessary to measure the quality of the pupil material with which the teacher is working. The achievement quotient is a useful device for doing this when standardized tests are used. In the case of measurement by means of written examinations and estimates of achievement, one should attempt to approximate the achievement quotient.

*Numerical score:*⁹ approximately the same as the achievement quotient (A. Q.); median score, 100; maximum score, 150; minimum score, 50.¹⁰

⁹The numerical score for this and other divisions of the plan proposed for measuring teaching efficiency implies a scheme of weighting the various divisions. The numerical score for achievements of pupils is for a relatively complete measurement of all achievements. In case only a few of the achievements of the pupils are measured this division should be given less weight.

¹⁰Unusually high or unusually low achievement quotients probably involve relatively large errors. For this reason upper and lower limits are specified.

II. PERSONAL AND PROFESSIONAL TRAITS OF THE TEACHER

The teacher is to be rated by means of a man-to-man comparison scale for each of the following four groups of traits: (a) interest in school work, particularly classroom instruction, (b) skill in the mechanics of managing a class, (c) quality of growth and keeping up-to-date, (d) personal and social qualities.

Numerical score: the rating for each group of qualities is to be on a scale for which the maximum score is 38. (See page 10.) The numerical score for this division is to be the sum of the four ratings divided by two. The maximum score according to this plan would be 76, median score 44, and minimum score 12.

III. GENERAL INTELLIGENCE OF THE TEACHER

For measuring the general intelligence of a teacher a scale suitable for adults should be used. The Otis Group Intelligence Test, Advanced Examination, or the Army Alpha is suggested.

Numerical score: A teacher's score on the intelligence test should be judged with respect to norms for teachers. In translating the test score into the numerical score for this rating scale, a test score equal to the average score for teachers should be taken as equivalent to 25. The maximum numerical score for intelligence should be 40 and the minimum 10.¹¹

IV. TEACHING EXPERIENCE

Numerical score: beginning teachers, 0; 4 points additional for each year of experience up to six years. Beyond six years it has been found that experience does not seem to be a potent factor in contributing to the success of the teacher.

V. ACADEMIC AND PROFESSIONAL TRAINING OF TEACHERS

Numerical score: (a) Teachers in the elementary school: completion of the eighth grade, a score of 0; completion of the twelfth grade, a score of 10; an additional two points for each six weeks of attendance at a college or a normal school. (b) Teachers in high school: completion of the twelfth grade, a score of 0; an additional five points for each year of college or normal school work.

¹¹It is likely that different norms should be used in evaluating the general intelligence of elementary-school teachers, and of high-school teachers.

COMPOSITE PLAN FOR RATING TEACHERS

QUALITIES MEASURED	SCHEME OF NUMERICAL RATING			RATING ASSIGNED TEACHER			
	Minimum	Median	Maximum				
I. Achievements of pupils..... (Similar to A. Q.)	50	100	150				
II. Personality of teacher..... (Man-to-man scale)	12	44	76				
III. Intelligence of teacher..... (Standardized scale)	10		40				
IV. Experience of teacher.....	0	4 points for each year of experience up to 6 years.	24				
V. Training of teacher..... (Elementary school)	0	2 points for each six weeks of attendance at a college or normal school.					
Training of teacher..... (High school)	0	5 points for each year of work at a col- lege or normal school.					
Composite Rating.....	Total of above items.....						

ANNOTATED BIBLIOGRAPHY ON MEASUREMENT OF
TEACHING EFFICIENCY

I. Studies of the qualities of teachers.

ANDERSON, W. N. "The selection of teachers," *Educational Administration and Supervision*, 3:83-90, February, 1917.

A list of qualities rated by presidents of school boards and another list rated by superintendents is included in this reference. Comparison is made between these lists and another list derived from an investigation of "Why teachers fail."

BIRD, GRACE E. "Pupil's estimate of teachers," *Journal of Educational Psychology*, 8:35-40, January, 1917.

A graph showing a comparison between the amount of quality desired by the high-school boy, the high-school girl, and the normal-school girl is given. A table of results is also included.

BOOK, W. F. "The high school teacher from the pupil's point of view," *Pedagogical Seminary*, 12:239-88, September, 1905.

This is an inquiry of pupils' opinions to discover traits in teachers that best aid pupil progress.

BOYCE, A. C. "Qualities of merit in secondary school teachers," *Journal of Educational Psychology*, 3:144-57, March, 1912.

This is a careful study of each particular quality in its relation to general teacher-merit. Tables of correlation are included.

BUELLESFIELD, HENRY. "Causes of failures among teachers," *Educational Administration and Supervision*, 1:439-45, September, 1915.

A list of qualities in order of frequency is included.

CLAPP, F. L. "Scholarship in relation to teaching efficiency." *School Review Monographs*, No. 6. Chicago: University of Chicago, 1915, p. 64-70.

COFFMAN, L. D. "The rating of teachers in service." *School Review Monographs*, No. 5. Chicago: University of Chicago, 1914, p. 13-24.

The correlations of Ruediger and Strayer, Boyce, Clapp, Littler, and Moses are included.

COLVIN, STEPHEN S. "The most common faults of beginning high-school teachers," *School and Society*, 7:451-59, April, 1918.

COURTIS, STUART A. "Standards of teaching ability," *Educational Review*, 62:183-86, October, 1921.

This is a short article on the present situation of teacher measurement.

JOHNSTON, JOSEPH H. "Teacher rating in large cities," *School Review*, 24:641-47, November, 1916.

KNIGHT, F. B. "Qualities related to success in teaching." *Teachers College Contributions to Education* No. 120. New York: Teachers College, Columbia University, 1922, 67 p.

371
K749

KNIGHT, F. B. "Qualities related to success in elementary-school teaching," *Journal of Educational Research*, 5:207-16, March, 1922.

An attempt was made in this article to contribute statistically dependable facts as to teacher selection.

KRATZ, H. E. *Studies and Observations in the Schoolroom*. Chicago: Educational Publishing Company, 1907, Chapter 5.

This is an investigation to discover the teaching qualities of teachers of the elementary school from the pupils' point of view.

LITTLER, SHERMAN. "Causes of failure among elementary-school teachers," *School and Home Education*, 33:255-56, March, 1914.

The list of qualities found by Littler is compared with those included in the study of Ruediger and Strayer.

MOSES, CLEDA. "Why high-school teachers fail," *School and Home Education*, 33:166-69, January, 1914.

The qualities of this study are also compared with those of Ruediger and Strayer.

OSBURN, J. W. "Personal characteristics of the teacher," *Educational Administration and Supervision*, 6:74-86, February, 1920.

This is a very comprehensive study which reviews the qualities of teacher merit discovered by the more important studies to date. It also includes a study of teaching qualities from the normal-school and college point of view.

PAYNE, E. GEORGE. "Scholarship and success in teaching," *Journal of Educational Psychology*, 9:217-19, April, 1918.

RUEDIGER, WILLIAM C. and STRAYER, GEORGE D. "The quality of merit in teachers," *Journal of Educational Psychology*, 1:272-78, May, 1910.

This is a preliminary study of teacher merit based upon fourteen qualities.

SEARS, J. B. "The measurement of teaching efficiency," *Journal of Educational Research*, 4:81-94, September, 1921.

A brief history, the theoretical and the practical aspects of teacher measurement, and an outline of the next step in teacher rating is included in this review.

WHITNEY, E. L. "The analysis of teaching functions," *Journal of Educational Research*, 7:297-308, April, 1923.

Partial correlations of various elements in teaching success are given.

II. Rating Scales for Teachers.

ABEL, E. L. "A practical teacher rating card," *The American Schoolmaster*, 14:256-58, September, 1921.

A score card is included in which an attempt was made to place overlapping qualities of a type under a single head.

ADAMS, W. C. T. "Two teacher rating cards," *The American School Board Journal*, 59:30, 101, December, 1919.

ADAMS, W. C. T. "Superintendent's Rating of Teachers," *Journal of Education*, 90:288-89, September, 1919.

A score card is included which divides teaching power into four elements and weights each equally, in percent.

"Basic principles in the making of a salary schedule for teachers," *The American School Board Journal*, 56:26-27, 83, March, 1918.

A score card of five main heads and fifty-five sub-heads is included. The aim of rating the teacher is to formulate the salary schedule.

✓ BOYCE, A. C. "A method of guiding and controlling the judging of teaching efficiency," *School Review Monographs*, No. 6. Chicago: University of Chicago, 1915, p. 71-82.

This article contains Boyce's Rating Scale, a lengthy discussion on methods of use, and the correlation of each of the forty-five qualities with general merit. A rank order is also given.

BOYCE, A. C. "Methods of measuring teachers' efficiency," *Fourteenth Yearbook of the National Society for the Study of Education*, Part II. Bloomington, Illinois: Public School Publishing Company, 1915, 82p.

This article by Boyce is the most complete discussion of his rating scale. It includes the scale, its method of use, method of reduction to a numerical rating, correlations of each of the qualities with general merit, and the rank order of the qualities.

BOYCE, A. C. "Qualities of Merit in Secondary School Teachers," *Journal of Educational Psychology*, 3:144-57, March, 1912.

BRACKEN, JOHN L. "The Duluth system of rating teachers," *Elementary School Journal*, 23:110-19, October, 1922.

Frequent supervision of teachers by supervisors is recommended. A score card is included.

BRADLEY, J. H. "A study of the relative importance of the qualities of a teacher and her teaching in their relation to general merit," *Educational Administration and Supervision*, 4:358-63, September, 1918.

Contains a rating scheme similar to that of Elliott's.

CARRIGAN, ROSE A. "The rating of teachers on the basis of supervisory visitation," *The Journal of Educational Method*, 2:48-55, October, 1922.

This score card contains rubrics which can be observed by a supervisor during a single visit. An attempt to avoid overlapping was made.

CLARK, M. G. "Sioux City, Iowa, teachers service standard," *School and Home Education*, 40:167-68, May-June, 1921.

A score card of five main heads and twenty-three rubrics.

CLARK, R. C. "A scale for measuring teachers," *American School Board Journal*, 62:39-40, February, 1921.

CONNOR, WILLIAM L. "A new method of rating teachers," *Journal of Educational Research*, 1:338-58, May, 1920.

COOK, WILLIAM A. "Uniform standards of judging teachers in South Dakota," *Educational Administration and Supervision*, 7:1-11, January, 1921.

The Dakota Rating Scales are included.

CRANOR, KATHERINE T. "A self-scoring card for supervisors as an aid to efficiency in school work," *Educational Administration and Supervision*, 7:91-102, February, 1922.

A self-scoring scheme for supervisors is included.

ELLIOTT, E. C. "How shall the merit of teachers be tested and recorded?" *Educational Administration and Supervision*, 1:291-99, May, 1915.

Principles underlying the formation of score cards are suggested.

FICHANDLER, A. "A study in self-appraisal," *School and Society*, 4:1000-02, December, 1916.

This study shows an agreement between the self-rating of the teachers and the independent rating of the principal. The criticisms of Taylor and Myers should be read with this article.

FOSTER, F. M. "A score card for rural teachers," *School and Society*, 12:131-32, August, 1920.

GRAY, WILLIAM S. "Rating scales, self-analysis, and the improvement of teaching," *School Review*, 29:49-57, January, 1921.

This is a general discussion which contains a rating scale of ten general traits included in Boyce's Rating Scale.

HERVEY, H. D. "The rating of teachers," *Journal of Education*, 93:319-20, March, 1921.

This article contains criteria for a successful merit system as well as an argument against present day systems of rating and their use.

HESS, ADAH H. "Teacher rating as a means of improving home-economics teachers in service," *The Journal of Home Economics*, 14:85-90, February, 1922.

Two rating schemes are included, namely the teacher self-rating scheme and the supervisors' rating scheme for home-economics teachers.

HICKMAN, JOSEPH. "A measuring scale for teachers in service," *American School Board Journal*, 52:43-44, April, 1916.

The score card included contains a rating scale of five general heads and eighteen sub-heads.

HILL, C. W. "The efficiency ratings of teachers," *Elementary School Journal*, 21:438-43, February, 1921.

JOHNSTON, JOSEPH H. "Scientific supervision of teaching," *School and Society*, 5:181-88, February, 1917.

KENT, RAYMOND A. "What should teacher rating schemes seek to measure?" *Journal of Educational Research*, 2:802-07, December, 1920.

The Duluth rating scheme is described.

✓ KNIGHT, F. B. "The effect of the 'acquaintance factor' upon personal judgments," *Journal of Educational Psychology*, 14:129-42, March, 1923.

A score card, diagrams, curves, tables, and charts showing various results of the "acquaintance factor" are included.

LANDSITTEL, F. C. "Evaluation of merit in high-school teachers," *School and Society*, 6:774-80, December, 1917.

LANDSITTEL, F. C. "A score card method of teacher rating," *Educational Administration and Supervision*, 4:297-309, June, 1918.

LAPORTE, WILLIAM RALPH. "A system of personality ratings for prospective physical-training teachers," *American Physical Education Review*, 27:23-24, January, 1922.

MADDOCK, WILLIAM. "Teacher rating." *Yearbook of the National League of Teachers' Association*, 1922-23, p. 36-66.

This paper includes a statement on the general trend of teacher rating, present day status, criticisms, criteria for the formation of a scale with an explanation of items, its application to a salary schedule, and throughout emphasizes the job of the supervisor.

MEYERS, GARRY C. "Teachers' rating," *School and Society*, 5:322-23, March, 1917.

This is a criticism of Fichandler's self-rating scheme and suggests a scheme of "teacher self-rating of one another."

MORTON, R. L. "Qualities of merit in secondary-school teachers," Educational Administration and Supervision, 5:225-39, May-June, 1919.

A score card and a comparison of city and rural teachers is included in this report.

RUGG, H. O. "Rating scales for pupils' dynamic qualities: standardized methods of judging human character," School Review, 28:337-49, May, 1920.

This article deals mainly with pupil rating. Its value here is the study of dynamic qualities.

SCHOWALTER, B. R. "A score card for rural teachers," School and Society, 12:200, September, 1920.

TAYLOR, JOSEPH S. "Measurement of educational efficiency," Educational Review, 44:348-67, November, 1912.

Contains a teacher rating card especially devised for kindergarten and another for elementary and high-school teachers. A school summary blank is also included.

"Teacher personality," American Physical Education Review, 26:50, January, 1921.

A detailed rating card of personality is given.

"Teacher rating card," Elementary School Journal, 20:723-24, June, 1920.

The score card included is the Omaha Rating Card.

VAN SICKLE, J. H., WHYTHE, JOHN, and DEFFENBAUGH, W. S. "Public education in the cities of the United States." U. S. Bureau of Education Bulletin, No. 48. Washington, 1918, 46 p. (Advanced Sheets from the Biennial Survey of Education in the United States 1916-1918.)

Four score cards are given.

WAGNER, C. A. "The construction of a teacher-rating scale," Elementary School Journal, 21:361-66, January, 1921.

Suggestion is used as the basis for teacher rating.

WITHAM, E. C. "School and teacher measurement," Journal of Educational Psychology, 5:267-78, May, 1914.

The rating scale included contains forty-six items among which curriculum studies are included.

III. The Man-to-Man Comparison Scales.

RUGG, H. O. "Self-improvement of teachers through self-rating: a new scale for rating teachers' efficiency," *Elementary School Journal*, 20:670-84, May, 1920.

The self-rating scale, Form I and Form II, instructions for use, and a concise description of the Man-to-Man rating scale are included in this article.

RUGG, H. O. "Is the rating of human character practical?" *Journal of Educational Psychology*, 12:425-38, 485-501; November and December, 1921; 13:30-42, 81-93; January and February, 1922.

These articles make up a very complete study of teacher rating based on studies carried on in the various camps during the World War. The Army Rating Scale, as well as Rugg's suggested scale, with instructions for use are included in the series of articles.

WOOLLEY, PAUL V. "The use of a scale for judging manual arts teachers," *Manual Training Magazine*, 23:5-8, July, 1921.

This is Rugg's Self-Rating Scheme adjusted to the needs of manual-arts teachers.

IV. Teacher Rating by Means of Standardized Tests and the Accomplishment Quotient.

ALMACK, JOHN C. "Keeping up in teaching," *The American School Board Journal*, 59:27-30, November, 1919.

A statement in this article says that educational tests are doing much to eliminate the uncertainty of teaching results. A rating scale of twelve points is included.

BLISS, W. B. "How much mental ability does a teacher need?" *Journal of Educational Research*, 6:33-41, June, 1922.

A comparison is made between teaching success and mental ability.

DOUGLAS, H. R. "Some uses and limitations of the standard educational test," *Educational Administration and Supervision* 5:475-90, December, 1919.

Douglas believes that the progress of pupils as measured by tests is a measure of a teacher's efficiency. Limitations of tests are also discussed.

JENKINS, ALBION U. "The measurement of teaching efficiency by means of standardizing tests," *The First Yearbook of the Department of Elementary School Principals*. Washington: National Education Association of the United States, 1922, p. 25-34.

A study showing the need of the A. Q. as a measure of teacher efficiency is included. Tables of results are also given.

JONES, EDWARD S. "A suggestion for teacher measurement," *School and Society*, 6:321-22, September, 1917.

A mental test is proposed as a means of measuring teacher efficiency.

KIRKPATRICK, E. A. "Intelligence tests in Massachusetts State Normal Schools," *School and Society*, 15:55-60, January, 1922.

Thurstone Intelligence Tests are used as a means of determining teacher efficiency.

MONROE, WALTER S. "An Introduction to the Theory of Educational Measurement." Boston: Houghton Mifflin Company, 1923, p. 272-74.

Measurement by standardized tests is suggested but due regard must be given to their limitations, which are mentioned in the article.

STEBBINS, R. C. "Accomplishment quotients as an aid in diagnosis," *The First Yearbook of the Department of Elementary School Principals*. Washington: National Education Association, 1922, p. 34-44.

Tables and graphs of the results of the study are included.

V. Miscellaneous Articles on Teacher Rating

BYRNE, LEE. "A method of equalizing the rating of teachers," *Journal of Educational Research*, 4:102-08, September, 1921.

A statistical treatment for equalizing ratings.

COFFMAN, L. D. "Committee on rating, placing and promotion of teachers," *School Review Monographs*, No. 6. Chicago: University of Chicago, 1915, p. 61-63.

Five recommendations on measurement of teacher merit are included in this brief committee report.

MILLER, GEORGE F. "Rating a teaching position," *American School Board Journal*, 58:35-36, February, 1919.

A teacher's point of view.

MORRISON, J. CAYCE. "Methods of improving classroom instruction used by helping-teachers and supervising-principals of New Jersey," *Elementary School Journal*, 20:208-16, November, 1919.

This is an inquiry by questionnaire to discover how supervisors rate and aid teachers.

PITTENGER, B. F. "Problems of teacher measurement," *Journal of Educational Psychology*, 8:103-10, February, 1917.

This article points out the limitations of score cards and the possibility of error in using them.

WAGNER, C. A. "Reducing the difficulties in rating and grading of teachers," American School Board Journal, 59:54-55, November, 1919.

This article suggests that Boyce's forty-five points be so divided that the person rating a particular point is the person knowing most about it.

WAGNER, C. A. "Some difficulties in rating and grading teachers," American School Board Journal, 59:28, September, 1919.

This is an optimistic treatment of rating.

WEBB, L. W. "One element to be considered in measuring effective teaching," School and Society, 13:206-09, February, 1921.

A questionnaire on method of studying is included.

THE BOARD OF THE
JAN 25 1927

UNIVERSITY OF ILLINOIS



UNIVERSITY OF ILLINOIS-URBANA



3 0112 084223004